STAFF SUMMARY SHEET

	TO	ACTION	SIGNATURE *(Surname)*, GRADE AND DATE		TO	ACTION	SIGNATURE *(Surname)*, GRADE AND DATE
1	DFCS	sign	David A. Gibson, Col 7 Mar 12 Col Gibson	6			
2	DFER	approve	Brent A Ritter, Col 8 Mar 2012	7			
3	DFCS	action	8 Mar 12 Dr. Fulton	8			
4				9			
5				10			

SURNAME OF ACTION OFFICER AND GRADE	SYMBOL	PHONE	TYPIST'S INITIALS	SUSPENSE DATE
Dr. Steven Fulton	DFCS	333-4126		

SUBJECT
Clearance for Material for Public Release USAFA-DF-PA-92

DATE
20120224

SUMMARY

1. PURPOSE. To provide security and policy review on the document at Tab 1 prior to release to the public.

2. BACKGROUND.
Authors: C1C Jacob Blasbalg, C1C Ryan Cooney, Dr. Steven Fulton

Title: Defining and Exposing Privacy Issues with Social Media

Circle one: Abstract Tech Report Journal Article Speech [✓ Paper] Presentation Poster

Thesis/Dissertation Book Other: ____________

Check all that apply (For Communications Purposes):

[] CRADA (Cooperative Research and Development Agreement) exists

[] Photo/ Video Opportunities [] STEM-outreach Related [] New Invention/ Discovery/ Patent

Description: Research Paper which focuses on student knowledge of privacy issues.

Release Information:

Previous Clearance information: (If applicable): N/A

Recommended Distribution Statement: (Distribution A, Approved for public release, distribution unlimited.)

3. DISCUSSION. This paper is simultaneously being reviewed by the National Security Agency pre-publication review team

4. VIEWS OF OTHERS. The Department Research Director has reviewed this paper and recommends it for public release.

5. RECOMMENDATION. Sign coord block above indicating document is suitable for public release. Suitability is based solely o the document being unclassified, not jeopardizing DoD interest, and accurately protraying official policy.

Steven P Fulton
Assistant Professor, USAFA/DFCS

1 Tab
Paper for approval

Defining and Exposing Privacy Issues with Social Media

Jacob Blasbalg, Ryan Cooney, Steven Fulton, United States Air Force Academy

ABSTRACT

In the growing world of social media, privacy concerns have grown out of an explosion of usage. Millions take advantage of the convenience found in websites such as Twitter and Facebook. One concern about such sites is that there is limited understanding of the specific privacy issues associated with them. Recent changes to Google's privacy rules are one such example. Many factors however, lead to the ignorance of users and therefore commonplace exploitation of the information which they knowingly or unknowingly share. This paper looks at freshmen university students who use social networking sites in an attempt to understand what expectations they have of privacy on social networking sites. Using a two phased process, the study first analyzes how well students understand privacy associated with their social networking posting and then identifies information from our students which may be of a personal nature that is publically available on social networking websites.

Index terms: Social Networking, Privacy, Student Privacy

I. INTRODUCTION

The increase in use of online social networks has inevitably led to more personal information being posted online. High school and college students, sports teams, and organizations often rely on Facebook or similar social networking sites, expecting its members or others to have membership to the site. Other sites such as MySpace, Twitter, and YouTube, like Facebook, can penetrate users' daily lives and lead to unintended consequences which threaten privacy of users and are gradually changing the relationship between public and private information [1]. Inadvertent disclosures of information are common, possibly as a result of the young age of most of social media's user base. As the number of user accounts grow, so do privacy concerns. Simple Google searches can yield incredible amounts of information, leading to both monetary and identity theft, highlighting the ignorance of many users[1].

II. LEGAL CLASSIFICATION OF PRIVACY

United States Federal Law is deeply embedded within the Fourth Amendment of the Constitution, which protects against unwarranted search and seizure. Privacy Law has become convoluted with the Internet however. The courts have generally agreed that when you use the Internet, it is the equivalent of leaving your home therefore the right to privacy is quickly lowered [1].

The Privacy Act has implications in the social media world. Congress passed the Privacy Act in 1974 focusing on sustaining traditional major privacy principles including the right to understand what records are being kept by the government on individuals [2]. Shortly after the passing of the act, military recruiters began creating a *marketing and recruitment database* full of "ethnicity, phone numbers, e–mail addresses, intended fields of study and extracurricular activities" collected from publically available social media sources[1]. This behavior would have been completely legal had it not been for the Privacy Act, as the system had been created prior to notifying the public [3]. The issue, however, was that the government was collecting the information. The Privacy Act only applies to the fact that government organizations were gathering the data. If, instead of military recruiters, it was private businesses collecting and keeping the information, the Privacy Act would have no applicability.

Barnes [3] states that in the eyes of the law, children 13 and under are considered minors and have legal protections which are too strong for many social sites to deal with. MySpace, one of the founding websites of the social media revolution, asked its users for age in order to attempt to block anyone from under the age of 14 from joining. However, since no age verification system was adopted it raised criticism from officials, such as Massachusetts Attorney General Tom Reilly, who brought attention to privacy problems when he asked both for an age verification system and that the site increases the staff responsible for reviewing pages for improper posts.

In reality, a user over the age of 13 that unknowingly posts information about themselves on the internet is not covered under the law unless someone uses it with the intent to defraud the user and obtain something of value

ISBN 1-933510-95-1/$15.00 □ 2012 CISSE

according to the Computer Fraud and Abuse Act [4]. Information on social networking sites have been used by the government itself for informational purposes and a growing number of employers and educational institutions use information found on these sites to build a character profile.

Privacy laws in the United States cannot seem to keep up with the technological advances seen in recent years. As a result, computers have free reign to collect as much information as possible. Every time a consumer does anything from writing a check to calling an 800 number, it leaves a data trail which is permanently stored in computerized databanks [5].

One of the largest concerns today is the lack of knowledge in the privacy dealings of how these sites conduct themselves. Long user agreements are inevitably ignored by the user or simply agreed upon because the jargon and complexity of it is beyond anyone without a law degree.

Interestingly, users are aware of the problems that exist with social networks. In a recent college survey, students were asked to respond to the statement: "*Facebook respects my privacy*." The student responses were neutral with a slight tendency to disagree with the statement. Similarly when asked "*In mediated environments like Facebook, my personal privacy is made public*." The student responses were again neutral. While students may be aware that Facebook has limited privacy, many continue to act as if messages posted are actually private [3].

III. How Privacy is Perceived

The perceived anonymity of social networking sites has caused a recent phenomenon of publishing social fantasy. The same urge to live in a fantasy atmosphere made applications like *Second Life* and *The Sims* record breaking computer hits is having its effect in social media. While some users want to keep their information private, especially embarrassing activity, an increasing amount of innocent younger users fantasize and embellish. For example, some young girls have blogged about weekends of drinking when in reality they did not do anything at all [3]. This behavior can be devastating if colleges find archives of it years later, or employers find it when sifting through applications. It is clear that while using these sites if your end goal is to publicize yourself, privacy is not your main concern.

Many users, as demonstrated in the marketing community, are willing to give up information about themselves if there is a perceived benefit. It seems that consumers are also likely to provide personal information if they believe that have control of the information, the information appears to be relevant and if the information is likely to create valid inferences about the preferences of the users. [6]. In the case of social media, users absolutely feel they have control over their information as evident by how much of it they voluntarily write on their postings. The problem lies with how such information is used however. Many MySpace and Facebook users do not understand that if their profile is set to "private" their information can still be used and saved by the owners of the web site. This violates the idea of *perceived privacy*, or the impression that the consumers have that the collection, access, use and disclosure of their private personal information is consistent with their beliefs regarding the way that information is being used [6]."

The complexity of how information is used may escape most consumers. For example, Insurance companies can purchase this information on a potential insured to determine if the insured may be a greater risk to the company. The result ends up being a fairly complete picture of a person's lifestyle[5] . This type of information release can potentially come from something as simple as buying goods from a grocery store with a credit card or more advanced as buying information posted in a social networking website.

IV. Privacy and the Role of the Computer

The computer has changed the concept of privacy on almost every level. Consumers that shop online and in a store face almost entirely sets of rules of what can and cannot be done with their consumer data. In a computer world with no more "sensory" borders, rules such as "if I can see you, you can see me" are no longer relevant [7]. Furthermore, computer designers for many websites find that privacy is not their problem. They use excuses like their projects are "only prototypes," or "firewalls will handle it," or that the problem is in the hands of the lawyers and politicians and not the designers themselves [7]. This can be a result of the complexity of the issue, or a wild-west mentality of a new internet culture that accepts information runs free.

Consumers and social media users deal with data collection that is invisible. Card swiping and form signing are now replaced with practically nothing [3]. Data collection done online is the equivalent of watching a person walk around a store and keep track of every object they look at. These websites will inevitably talk together to form a user profile. Sites such as Amazon, Google, and Facebook accounts will bring together information on an unprecedented scale. Users have little say in the matter. In order to use these services, you consent to this type of information sharing. This type of profiling already exists. For example, Gmail reads messages to recommend products and services to be posted in advertisements. Any day now, Google plans to change the way it uses data to tie its different products (web searches, gmail, Google+) so that information can be gathered across the different applications for use in advertisement. Such a change is unprecedented and currently being challenged by several US States[8]

V. Social Networking Websites

Despite the risks to privacy, social networking websites provide their users with the convenience of connecting with friends, classmates, colleagues, and even family. Although formally defining what constitutes a social networking website may seem trivial due to their common use today, doing so provides a foundation for comparing different sites and for tracking the history and evolution of social networking.

In the days of dialup, companies such as America Online took advantage of the technologically inept by charging for a dialup connection and a pretty browser interface. Most do not consider America Online to be social networking because users could not browse their buddy's buddy lists. However users were able to create very limited profiles, add people they wished to communicate with using Instant Messenger to a Buddy List, and browse the profiles of others. According to the same article above, the first true social networking website was SixDegrees.com, a play on the supposed "Six Degrees of Separation" between any two people on the globe. SixDegrees was launched in 1997. Like America Online, the primary purpose of SixDegrees was to enable users to communicate through messages [9]. SixDegrees failed mostly due to disinterest at the time. However, the service set a precedent for social networking. Almost 15 years later, the top three most commonly used social networking sites are Facebook, Twitter, and LinkedIn[10].

VI. Search Engines

In addition to social networking sites, search engines pose new issues to privacy. As previously mentioned, something as harmless as swiping a credit card can leave a data trail. Public figures have even more information to worry about online, considering people write biographies which include many details from their lives. One of the most infamous cases of using simple search methods to gain unauthorized access to private information occurred during the last presidential campaign. Vice Presidential candidate Sarah Palin's email account was hacked, and her personal emails were posted throughout the internet.

The perpetrator admitted, "It took seriously 45 minutes on Wikipedia [sic] and Google to find the information. Birthday? 15 seconds on Wikipedia, zip code? Well she had always been from Wasilla... The second was somewhat harder, the question was 'where did you meet your spouse?'... I found out later through more research that they met at high school, so I did variations of that... eventually hit on 'Wasilla high.' [2]"

VII. Final Background Thoughts

Social networking, search engines, and storing personal information online in general have been accepted worldwide due to the benefits they provide. Social networking provides even more communication in an information-demanding age, allowing users to interact across great distances. Search engines allow people to find nearly anything as fast as they can type and read. However, negative aspects almost always emerge as people begin exploiting new technology. People in general are not as safe as they think they are online, and inadvertently or intentionally allow private information to be public. The only way to combat risks such as identity theft is to limit the amount of information available to the public.

VIII. Research Goal and Approach

The intent of this project was to expose previously ignored issues regarding publically available information on social media websites. We used freshmen at our university. Since the entire freshman is required to take a core computer science course, it gave us the chance to issue a mass survey in a formal environment.

In the first phase of this study, students were asked questions regarding their use of social networks as well as several questions regarding their perceived privacy when using social network sites. For example, students were asked to respond to questions such as "I am concerned about privacy issues in general with respect to social networking websites" and "I have intentionally posted information on the internet about myself or others to the public that I would not want others to see" to measure student education on social network privacy settings, as well as how secure they think they are. We surveyed a sample of n=98 out of a class size of about 1000.

The second phase of the study attempted to verify the findings of phase 1. Using the most popular social networking web sites, we identified 82 students who publically information identifying themselves as freshmen from our institution. We did not know if these students were from the same population as our original students. According to the original questionnaire, students predominantly use Facebook for social networking purposes. Using Facebook as a starting point, names of students from the class of 2015 were identified by using standard search requests. From that point, we attempted to identify a series of information regarding that student. Facebook, Google+, and Twitter were all used as resources, being the only predominantly used social media sites used by students as found by the survey. No personal information was saved, and when privacy information was found (such as a phone number or address), the fact that the information was found was recorded but not the information itself. This was to ensure that privacy concerns were satisfied.

IX. Results

The questions asked in the initial survey as well as the results from these questions are outlined in Table 1: Questionnaire Results.

Table 1- Questionnaire Results

Question	Scale	Mean
If you [belong to any]social networks, do you intentionally set your privacy setting so that only Facebook Friends or equivalent can access your information?	0=No, 1=Yes	0.91
Have you ever successfully used a search engine to find personal or uniquely identifiable information about yourself or someone else	0=No, 1=Yes	0.41
If answered Yes to the above question,, rate your agreement with the following statement: Until I used search engines to look up uniquely identifiable information, I was unaware of how much information existed on-line about me or the other person I searched	1=Strongly Agree, 5=Strongly Disagree	2.84 (Standard Deviation = 1.05)
I am concerned with privacy issues in general with respect to social networking websites	1=Disagree, 2=Neither Agree nor Disagree, 3=Agree	2.56 (Standard Deviation = 0.79)
I believe that users own the data they post on social networking sites	1=Disagree, 2=Agree	1.33
I believe that posting information on a personal page within a social network represents a potential privacy issue	0=Disagree, 1=Agree	0.87
I believe that privacy issues, with respect to social media sites, are a major issue in our society	1=Disagree, 2=Neither Agree nor Disagree, 3=Agree	2.54 (Standard Deviation = 0.76)
In hindsight, I have intentionally posted information on the internet about myself or others to the public that I would not want others to see	0=Disagree, 1=Agree	0.22
I have accidentally posted information on the internet about myself or others to the public that I would not want others to see	0=Disagree, 1=Agree	0.35
I make sure to stay aware of all privacy setting changes on social media sites that I use	0=Disagree, 1=Agree	0.76
I check and update my privacy settings any time I hear that there have been changes to social media sites	0=Disagree, 1=Agree	0.62
I am not concerned about who sees content I have posted on my personal pages within social media sites	0=Disagree, 1=Agree	0.29
I am concerned about privacy issues in general with respect to social media sites	0=Disagree, 1=Agree	0.72
I believe that what I post on my personal pages represents a possible privacy issue	0=Disagree, 1=Agree	0.71

Overall, students appeared to be concerned with their privacy settings. Ninety-one percent of those surveyed agreed felt that they intentionally set their privacy settings so only those who knew them could identify them (i.e. Facebook friends or equivalent). Similarly, when asked if they were concerned with privacy when using social networking sites, most students admitted to being concerned. About forty one percent of those surveyed admitted to having used a search engine to successfully find personal or uniquely identifiable information about themselves or others. Additional findings included most students realize that when they post information on a social networking site, they may no longer own such information. A full thirty five percent of those surveyed admitted to having accidentally posted information on a social networking site either about themselves or others they felt they may not have wanted others to see or know. Twenty-two percent of the surveys admitted having, in hindsight, posted information about themselves or others which they may not have wanted others to see.

When looking at relationships between responses, some interesting but not surprising results were discovered. Those who believed that they intentionally set their privacy settings so that only 'friends' could see their postings were more likely to believe that information posted on a personal page within a social network represents a potential privacy issue r(96)=.23, p<.05 and were more likely to have intentionally posted information about themselves or others to the public which they would not, in hindsight, want others to see r(89)=.21, p<.10.

Not expectantly, those who were more concerned about privacy issues in general with respect to social networking sites were likely to believe that what they posted on personal pages could represent a possible privacy issue r(91)=.31 p<.01. Also, those who admitted, in hindsight, intentionally posting information about themselves or others were likely check and update their privacy settings anytime they hear that a social media's site had been updated r(91)=.20, p<.05.

The second phase of this study attempted to find information regarding individual cadets available publically to understand the relationship of information provided in the anonymous survey with the actions they perform. This sample was independent of the first study. This information was found without being identified as a 'Facebook friend' or equivalent. The goal was to find information on a representational sample of the class population (n=84). A breakdown of the information discovered is provided in Table 2: Publically Available Information on Students.

Table 2 - Publically Available Information on Students

Information Type	Percent of Students with Information Publically Available.
Occupation/Employer	52.38%
Personal Pictures	21.43%
Date of Birth	4.76%
Address for home of record	7.14%
State and town of home of record	29.76%
State and town of current residence	50.00%
Social Networks used	4.76%
Marital Status	14.29%

Eighty-two cadets were identifying using the most popular social media site identified in the questionnaire as being students of our institution. Out of those 82 cadets, more than half of them identified the state of residence and their employer. About 22 percent of those identified supplied pictures. Something that is of more concern, however, would be the use of full addresses and birthdates. While only 7.14% provided full addresses and only 4.76% provided a birthdate, these are serious privacy issues which could be cause for concern.

X. Conclusion

Overall, it seems that students are generally concerned about privacy and feel that they have taken steps needed to secure the information on social media sites. It is interesting that the students who have either posted information that they feel should have been private have found information regarding other students which should have been kept private is of concern. What is troubling, however, is that so much information is available on social networking sites, even with student awareness of the issue. The fact that names associated with ages and home addresses available (and in some rare cases birthdays) in public searches of social networking sites suggests that knowledge, alone, isn't enough.

One possible solution is to make students aware of how much information can exist about a person publically without knowing more than a single piece of information such as a school or home town. This could be done either in formal classrooms setting or through a public awareness campaign. At our institution, our computer science core class has a section on computer privacy which has a lab exercise which has students identify information available publically on the web about a fellow cadet. The goal of this exercise is to attempt to make students aware of the existence of such information.

Regardless of how much a student feels that they are aware of privacy issues associated with information posted to social media sites, it is clear that their actions do not follow their level of understanding. It is imperative that we make it clear not only how important it is that the sites be upfront in how they plan to use and store information posted on their web sites, but users of such sites must understand what the impact of posting information means to them as individuals. Until the public is clear how information may be used, personal privacy continues to be at risk.

XI. Future Work

Future work should be directed in one of two areas: Legislation to protect the rights of privacy of the public and the awareness of the public as to what data is being saved and how best to limit access to that data. In our study, students who had posted information that they later felt was inappropriate were much more likely to be concerned with privacy settings on their profiles as well as the use of information by the owners of the social networking sites. Such awareness can change attitudes towards how information may be posted on public web sites.

XII. References

1. Debatin, B., et al., *Facebook and online privacy: Attitudes, behaviors, and unintended consequences.* Journal of Computer Mediated Communication, 2009. **15**(1): p. 83-108.
2. Campanile, C., *Dem Pol's Son was 'Hacker'*, in *New York Post*September 19, 2008: New York.
3. Barnes, S.B., *A privacy paradox: Social networking in the United States.* First Monday, 2006. **11**(9).
4. 99th Congress of the United States, *Computer Fraud and Abuse Act*, U.S. Congress, Editor 1986.
5. Petersen, S.B., *Your Life as an Open Book: Has Technology Rendered Personal Privacy Virtually Obsolete.* Fed. Comm. LJ, 1995. **48**: p. 163.
6. Chellappa, R.K. and P.A. Pavlou, *Perceived information security, financial liability and consumer trust in electronic commerce transactions.* Logistics Information Management, 2002. **15**(5/6): p. 358-368.
7. Lahlou, S., M. Langheinrich, and C. Röcker, *Privacy and trust issues with invisible computers.* Communications of the ACM, 2005. **48**(3): p. 59-60.
8. Vijayan, J., *States Challenge Google Privacy Policy Change*, in *Computer World*2012: http://www.pcworld.com/article/250698/states_challenge_google_privacy_policy_change.html.
9. Boyd, D. and N. Ellison, *Social Network Sites: Definition, History, and Scholarship.* Journal of Computer-Mediated Communication, 2008. **13**: p. 210-230.
10. eBiz MBA: The eBusiness Knowledgebase. *The 15 Most Popular Social Networking Sites, Sept 2011.* 2011; Available from: http://www.ebizmba.com/articles/social-networking-websites.

ISBN 1-933510-95-1/$15.00 □ 2012 CISSE